

A Stieltjes Transform Approach for Analyzing the RLS Adaptive Filter

Ali Vakili and Babak Hassibi
Electrical Engineering Department
California Institute of Technology
Pasadena, California 91125
Email: avakili, hassibi@caltech.edu

Abstract—Although the RLS filter is well-known and various algorithms have been developed for its implementation, analyzing its performance when the regressors are random, as is often the case, has proven to be a formidable task. The reason is that the Riccati recursion, which propagates the error covariance matrix, becomes a random recursion. The existing results are approximations based on assumptions that are often not very realistic. In this paper we use ideas from the theory of large random matrices to find the asymptotic (in time) eigendistribution of the error covariance matrix of the RLS filter. Under the assumption of a large dimensional state vector (in most cases $n = 10 - 20$ is large enough to get quite accurate predictions) we find the asymptotic eigendistribution of the error covariance for temporally white regressors, shift structured regressors, and for the RLS filter with intermittent observations.

I. INTRODUCTION

Adaptive filtering theory is of great importance in many areas such as communications and control. Basically adaptive algorithms appear whenever we encounter time-variant systems with little a priori information about the underlying signals. Nevertheless, analyzing the steady-state and transient behaviors of adaptive filters remains a formidable task in many cases.

One of the most important families of adaptive filters is the so-called *recursive least squares* (RLS) filter. As its name suggests, the RLS filter is an optimal filter that finds an estimate of a weight vector that recursively solves a least-square-error minimization problem. Moreover, it often demonstrates much faster convergence in its learning curve compared to other classes of adaptive filters such as *least mean square* (LMS) filter. There are several variants of the RLS filter which are mathematically equivalent but differ considerably in terms of computational expense and ease of implementation. Due to these reasons, the RLS filter has received much attention in the literature. Although the idea of the RLS algorithm has its roots going back to Gauss, when the regressors are random, as is the case in many applications, the RLS filter has proven to be one of the most difficult ones to analyze. In fact, the so-called excess mean-square error (EMSE) of an RLS filter is not known except for approximations based on several unrealistic assumptions.

This work was supported in part by the National Science Foundation through grant CCF-0729203, by the David and Lucille Packard Foundation, by the Office of Naval Research through a MURI under contract no. N00014-08-1-0747, and by Caltech's Lee Center for Advanced Networking.

The mean-square performance, as well as the tracking performance, of the RLS filter under an ergodic approximation and for temporally white regressors can be found in the literature [1], [2], [3], [4]. The ergodic assumption essentially asks for $\mathbb{E}P_i$ being equal to $(\mathbb{E}P_i^{-1})^{-1}$ where P_i is the estimation error covariance matrix. In [2, p. 319] it is shown that this is not a very good approximation for a wide range of forgetting factors. In this paper, we introduce a new method for analyzing the RLS filter when the number of taps is relatively large which does not need the above assumptions. We will stay in the framework of the generalized RLS filter [2, p. 772]

$$\begin{cases} x_{i+1} &= x_i + u_i \\ y_i &= h_i x_i + v_i \end{cases} \quad (1)$$

$$\mathbb{E} \begin{bmatrix} u_i \\ v_i \end{bmatrix} \begin{bmatrix} u_j^T & v_j^T \end{bmatrix} = \begin{bmatrix} qI & 0 \\ 0 & r \end{bmatrix} \delta_{ij} \quad (2)$$

in which $x_i \in \mathcal{R}^n$, the state vector we want to estimate, undergoes a random walk and u_i and v_i denote the zero-mean process and measurement noises, respectively. y_i is the measured signal and h_i is the so-called regressor vector. The regressor vectors are time-variant and often random. The RLS recursive estimate of x_i can be expressed as

$$\hat{x}_{i+1} = \hat{x}_i + \frac{P_i h_i^T}{r + h_i P_i h_i^T} (y_i - h_i \hat{x}_i) \quad (3)$$

and the estimation error covariance, $P_i = \mathbb{E}(x_i - \hat{x}_i)(x_i - \hat{x}_i)^T$, satisfies the celebrated Riccati recursion

$$P_{i+1} = P_i - \frac{P_i h_i^T h_i P_i}{r + h_i P_i h_i^T} + qI, \quad P_0. \quad (4)$$

The above Riccati recursion is nonlinear and time-variant and, in general, does not converge. Furthermore, when the h_i are random, it is a random matrix recursion. When the h_i are stationary—an assumption we shall make—it may be expected that P_i will converge to a matrix-valued stationary random process. In this case, the stationary distribution of P_i , and its statistics, such as $\mathbb{E}P_i$ or $\mathbb{E}\text{tr}P_i$, will be of interest. This is what we intend to analyze in this paper.

We should also mention that the filter described by (1-4) is one of many variants of the RLS algorithm (such as those with a forgetting factor—see, e.g., [2], [3]). Although our analysis easily extends to those variants, in this paper we

shall focus on the basic filter (1-4). Finally it is also worth mentioning that the basic recursion (4) is closely related to that of the H^∞ adaptive filter [5], [6].

$$\hat{x}_{i+1} = \hat{x}_i + \frac{\mathbf{P}_i \mathbf{h}_i^T}{r + \mathbf{h}_i \mathbf{P}_i \mathbf{h}_i^T} (y_i - \mathbf{h}_i \hat{x}_i) \quad (5)$$

$$\mathbf{P}_{i+1} = \mathbf{P}_i - \frac{\mathbf{P}_i \mathbf{h}_i^T \mathbf{h}_i \mathbf{P}_i}{\frac{r}{1-\gamma^{-2}} + \mathbf{h}_i \mathbf{P}_i \mathbf{h}_i^T} + q\mathbf{I}, \quad (6)$$

$$\mathbf{P}_i^{-1} - \gamma^{-2} \mathbf{h}_i^T \mathbf{h}_i > 0 \quad (7)$$

Therefore our approach can readily be extended to analyze the H^∞ filter, too.

In Section II we briefly review those elements of the theory of large random matrices which we have used to develop our framework. In Section III we use these results to analyze the basic RLS filter with temporally uncorrelated regressors and we find the steady-state eigendistribution of the error covariance matrix. While this is certainly a new result in adaptive filtering theory, it also appears to be a new result in random matrix theory, as, to the best of our knowledge, the eigendistributions we obtain have never been encountered before. Moreover, the simulation results demonstrate a very close match between the theoretical and empirical curves. Section IV extends the results to include a more realistic structure for the regressor vectors by relaxing the uncorrelatedness condition. Finally in Section V we further generalize the results to include the case of intermittent measurements where the measurements $\{y_i\}$ are transmitted through a network and may be dropped with some probability. This problem is motivated by the recently growing interest in estimation and control over lossy networks. We show how this scenario can be handled easily through our method and how closely we can predict the empirical curve.

II. SOME RESULTS FROM RANDOM MATRIX THEORY

We will start by giving a brief overview of selected results from random matrix theory which have been used in our work. We will also establish our notation for the rest of the paper.

For an $n \times n$ random matrix \mathbf{M} , whose elements have some joint probability distribution, the cumulative eigenvalue distribution can be defined as,

$$F_{\mathbf{M}}(x) = \frac{1}{n} \sum_{l=1}^n \Pr \{ \lambda_l(\mathbf{M}) \leq x \}, \quad (8)$$

where $\lambda_l(\mathbf{M})$ is the l -th eigenvalue of \mathbf{M} . Now a probability density function can be assigned to the cumulative function defined as above. We will refer to this density function, $f_{\mathbf{M}}(\lambda)$, as the *eigendistribution* of the random matrix \mathbf{M} . Basically $f_{\mathbf{M}}(\lambda)$ is the marginal distribution of one randomly selected eigenvalue of \mathbf{M} .

In many cases, the eigendistribution of a random matrix demonstrates interesting properties. Although most of the existing results on the eigendistribution of fixed-dimensional random matrices are cumbersome and limited to special cases mostly involving Gaussian random matrices (or matrices

derived from Gaussian random matrices), when the matrix dimensions are allowed to grow, one can often obtain simple closed form results for the eigendistribution even without a Gaussian assumption through a number of very useful mathematical tools which are available for large random matrices. As a matter of fact, these results can usually be interpreted as universal laws which are independent of the underlying distributions—counterparts to the central limit theorem for scalar random variables. For example, for the Wigner matrix defined as [7],

$$\mathbf{W} = \frac{1}{\sqrt{2n}} (\mathbf{A} + \mathbf{A}^T) \quad (9)$$

where \mathbf{A} is an $n \times n$ random matrix having i.i.d. entries with zero mean, unit variance, and finite higher moments, the eigendistribution of \mathbf{W} as $n \rightarrow \infty$ converges to the semi-circle law [8],

$$f_{\mathbf{W}}(\lambda) = \begin{cases} \frac{1}{2\pi} \sqrt{4 - \lambda^2} & \text{when } -2 \leq \lambda \leq 2 \\ 0 & \text{otherwise} \end{cases}.$$

Clearly the above result does not depend on the actual distribution of the elements of \mathbf{A} . It should also be mentioned that in the case of random matrices, even for dimensions as low as $n = 10$, these asymptotic results are very close predictions of the eigendistribution (This is due to the fact that we are essentially averaging over n^2 entries.)

Many results in the theory of large random matrices are expressed in terms of some *transform* of the eigendistribution. There exist many transforms which may be used in different cases depending on the structure of the problem, e.g., R-transform, Mellin transform, η transform, etc. A complete list of these transforms and examples of their application can be found in [9]. The most important of them, the Stieltjes transform, was first used in the seminal work of Marcenko and Pastur [10]. Given a probability density function $f_{\lambda}(\lambda)$, its Stieltjes transform is defined as,

$$S_{\lambda}(z) = \mathbb{E} \left[\frac{1}{\lambda - z} \right]. \quad (10)$$

When $f_{\lambda}(\lambda)$ represents the eigendistribution of a random matrix, there exist equivalent definitions of the Stieltjes transform which turn out to be extremely useful. In fact, given a random matrix \mathbf{M} , the Stieltjes transform of its eigendistribution (which we will also denote as the Stieltjes transform of the matrix itself) can be expressed as,

$$S_{\mathbf{M}}(z) = -\frac{d}{dz} \mathbb{E} \frac{1}{n} \log \det(\mathbf{M} - z\mathbf{I}) \quad (11)$$

or equivalently,

$$S_{\mathbf{M}}(z) = \mathbb{E} \frac{1}{n} \text{tr} (\mathbf{M} - z\mathbf{I})^{-1}. \quad (12)$$

Having the Stieltjes transform of the distribution, we can retrieve the distribution uniquely through the well-known

inversion formula which basically asks for the imaginary part of the Stieltjes transform very close to the real line [11],

$$f_M(\lambda) = \lim_{\omega \rightarrow 0^+} \frac{1}{\pi} \text{Im}[S_M(\lambda + j\omega)]. \quad (13)$$

The moments of the distribution can also be obtained directly from the Stieltjes transform through power series expansions. When the Stieltjes transform is analytic everywhere outside a disc of a certain radius (which is the case when the eigendistribution has bounded support), the corresponding Laurent expansion in terms of z yields the moments $m_i = \mathbb{E}\lambda^i$, $i = 0, 1, \dots$

$$S_M(z) = - \sum_{i=0}^{\infty} m_i z^{-i-1}, \quad (14)$$

Similarly, when the Stieltjes transform is analytic everywhere inside a disc of a certain radius, the corresponding Taylor series expansion yields all the negative moments.

An important property of the Stieltjes transform which makes it very useful in random matrix theory is the self-averaging property which is stated in the following lemma [12].

Lemma 1 (Self-averaging property): Let A be an $n \times n$ positive semidefinite random matrix. As $n \rightarrow \infty$ the eigendistribution of A converges almost surely to its mean value, i.e.,

$$\lim_{n \rightarrow \infty} \frac{1}{n} \text{tr}(A - zI)^{-1} = S_A(z) \quad \text{a.s.}$$

(Note the absence of the expectation.) Then for any n -dimensional vector x with i.i.d. zero-mean, unit-variance elements which is independent of A , we have

$$\lim_{n \rightarrow \infty} \frac{1}{n} x^T (A - zI)x = S_A(z) \quad \text{a.s.} \quad (15)$$

The above lemma can be intuitively verified by comparing it to the definition (12).

III. RLS FILTER WITH TEMPORALLY UNCORRELATED REGRESSOR VECTORS

First let us consider the RLS adaptive filter assuming that the regressor vectors are spatially and temporally uncorrelated. Although this assumption is widely used in the literature for analysis purposes, it is often not too reasonable. (We will, in fact, generalize it to more realistic cases in the next section.) For the time being, any h_i is assumed to be comprised of i.i.d. entries that are further independent of all other $h_{j \neq i}$'s. This model results in the following Riccati recursion for the error covariance,

$$P_{i+1} = P_i - \frac{P_i h_i^T h_i P_i}{r + h_i P_i h_i^T} + qI, \quad P_0 \quad (16)$$

where h_i are independently drawn $n \times 1$ vectors having i.i.d. zero-mean, unit-variance entries with finite higher order moments, and q is a constant. Clearly P_i will not converge to any steady-state P . However, its probability distribution does converge to a steady state. The study of P_i is of great importance since, for example, $\mathbb{E}\text{tr}P_i$ is actually the mean

square error and determines the performance of our adaptive filter.

One also should note that due to the highly nonlinear nature of this recursion, all the moments of the eigendistribution are coupled and therefore one cannot find the moments of the distribution by simply averaging both sides of (16) as in the cases analyzed in [13]. However, using the theory of large random matrices, we can find the asymptotic behavior of the eigendistribution.

We will assume that the state vector's dimension is large. Since there is only one measurement at each time step, one can verify through (16) that the state vector update should have a variance which behaves like $q \propto \frac{1}{n}$, otherwise the recursion will diverge. Therefore we should consider (16) for $q = \frac{\beta}{n}$ where β is some finite constant.

Theorem 1: (RLS Riccati recursion with temporally independent regressors): Consider the random Riccati recursion in (16) where $n \rightarrow \infty$ and $q = \frac{\beta}{n}$. As $i \rightarrow \infty$, the Stieltjes transform of the eigendistribution of P_i converges to the steady-state Stieltjes transform, $S(z)$, which satisfies

$$\beta S(z) + c = -\log(r - z - z^2 S(z)). \quad (17)$$

where c is a constant to be determined.

Proof: We will use the definition of Stieltjes transform as in (11). By bringing the qI term from the RHS to the LHS of the equation and taking the Stieltjes transform from both sides, we will obtain

$$\begin{aligned} S_{i+1}(z + q) &= -\frac{1}{n} \mathbb{E} \frac{d}{dz} \log \det \left(P_i - \frac{P_i h_i^T h_i P_i}{r + h_i P_i h_i^T} - zI \right) \quad (18) \\ &= -\frac{1}{n} \mathbb{E} \frac{d}{dz} \log \det (P_i - zI) - \frac{1}{n} \mathbb{E} \frac{d}{dz} \log \det \end{aligned}$$

$$\left(I - (P_i - zI)^{-1} \frac{P_i h_i^T h_i P_i}{r + h_i P_i h_i^T} \right) \quad (19)$$

The first term in the LHS is just the $S_i(z)$. Since $\frac{1}{r + h_i P_i h_i^T}$ does not depend on z and $\det(I - AB) = \det(I - BA)$,

$$\begin{aligned} S_{i+1}(z + q) &= S_i(z) - \frac{1}{n} \mathbb{E} \frac{d}{dz} \log \det (r + h_i P_i h_i^T \\ &\quad - h_i P_i (P_i - zI)^{-1} P_i h_i^T) \quad (20) \end{aligned}$$

We will now invoke the self-averaging property of the Stieltjes transform (15) to obtain

$$\begin{aligned} h_i P_i (P_i - zI)^{-1} P_i h_i^T &= h_i P_i (P_i - zI)^{-1} (P_i - zI + zI) h_i^T \\ &= h_i P_i h_i^T + z h_i (P_i - zI + zI) (P_i - zI)^{-1} h_i^T \\ &= h_i P_i h_i^T + z h_i h_i^T + z^2 h_i (P_i - zI)^{-1} h_i^T \\ &\rightarrow h_i P_i h_i^T + z + z^2 S_i \quad (21) \end{aligned}$$

Replacing (21) into (20) we will obtain

$$S(z + q) - S(z) = -\frac{1}{n} \frac{d}{dz} \log(r - z - z^2 S(z)). \quad (22)$$

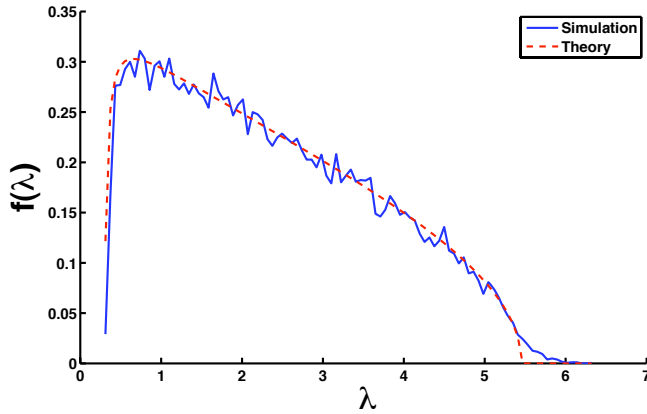


Fig. 1. Steady-state error covariance eigendistribution for the Riccati recursion of the RLS filter with temporally white regressors (13) for $n = 30$, $r = 2.5$, and $\beta = 0.5$

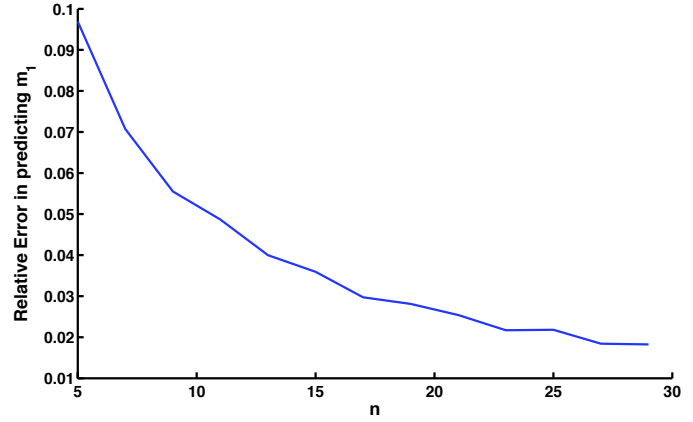


Fig. 2. Relative error in predicting $m_1 = \mathbb{E}\lambda = \frac{1}{n} \mathbb{E} \text{tr} P$ for the Riccati recursion of the RLS filter (13) for different values of n with $r = 2.5$ and $\beta = 2.5$

Finally, since $q = \frac{\beta}{n}$, by expanding $S\left(z + \frac{\beta}{n}\right)$ up to the first order term the LHS can be rewritten as,

$$S(z + q) - S(z) = \frac{\beta}{n} \frac{d}{dz} S(z), \quad (23)$$

and integrating both sides of the (22) yields (28). ■

Corollary 1: The constant c in (28) depends on the first moment of the eigendistribution as

$$c = -\log(r + m_1) \quad (24)$$

and it can be found numerically by insisting that the eigendistribution which results from $S(z)$ integrates to one, i.e. $\int \frac{f_\lambda(\lambda)}{\lambda - z} d\lambda = 1$.

Proof: We can use the Laurent power series expansion of the Stieltjes transform as in (14) to find an expression for c . In fact (28) can be rewritten as

$$-\frac{\beta}{z} - \frac{\beta m_1}{z^2} + \dots + c = -\log\left(r - z + z + m_1 + \frac{m_2}{z} + \dots\right) \quad (25)$$

which immediately results in (24). ■

Thus c depends on the first moment of the eigendistribution. This was expected due to the moments being all coupled; otherwise we could find all the moments readily through Laurent series expansion of $S(z)$. However, c can be numerically determined (with little effort) by simply insisting that the eigendistribution, which is directly related to the imaginary part of the Stieltjes transform, integrates to one.

Numerical methods such as the Newtonian step can be used to find the eigendistribution from (28). More specifically, one can break $S(z)$ into its real and imaginary parts and for every value of $z = \lambda$ solve a system of two coupled nonlinear equations to find the imaginary part of $S(z)$ which directly yields the eigendistribution according to (13). We have plotted the simulation results for $n = 30$, $\beta = .5$, and $r = 2.5$ in Figure 1. As can be seen, the plot shows

that the obtained theoretical expression closely predicts the actual distribution.

In order to show how the statistics of the steady state can be accurately found through our result, in Figure 2 we have plotted the error in predicting $m_1 = \mathbb{E}\lambda = \frac{1}{n} \mathbb{E} \text{tr} P$ for different values of n . We find m_1 through simulation and then compare it to the m_1 found by insisting that the eigendistribution should integrate to one. As can be seen in the plot, the relative error is less than 3 percents for $n > 20$ and less than 10 percents for n as small as 5.

As mentioned earlier, our result is certainly new in the adaptive filtering literature. However, we also believe it to be new in random matrix theory as (28), where the Stieltjes transform appears implicitly inside a logarithm function, is not something we have encountered before.

IV. GENERALIZATION TO THE SHIFT STRUCTURED REGRESSORS

So far we have considered regressor vectors which are spatially and temporally white. In this section we will generalize the results to include a more realistic model. We relax this assumption by showing that the results obtained earlier directly extend to the case where we have regressors with shift structure—a model that frequently arises in the literature when we have finite-impulse-response (FIR) adaptive filters (see, e.g., [3]). Thus, consider

$$h_i = [u_i \quad u_{i-1} \quad \dots \quad u_{i-n+1}] \quad (26)$$

where the u_i are drawn from a zero-mean, unit-variance white process. The following lemma proves to be useful in analyzing this case.

Lemma 2: For a given vector h_i and a randomly chosen orthogonal matrix Θ , there exists a matrix Δ such that

- 1) $h_i(\Theta + \Delta) = \|h_i\| e_1$
- 2) $(\Theta + \Delta)(\Theta + \Delta)^T = I$
- 3) $\mathbb{E} \text{tr}(\Delta \Delta^T) = O(\frac{1}{n})$

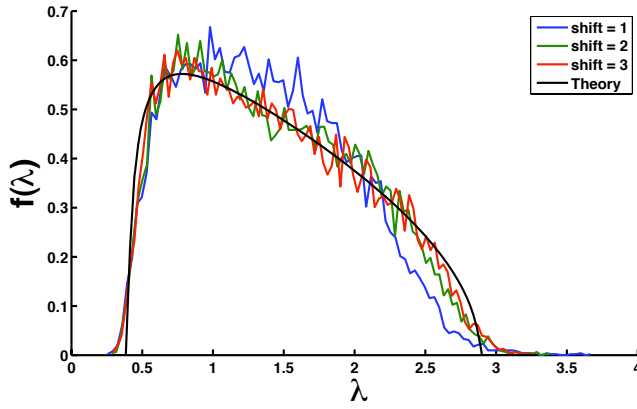


Fig. 3. Steady-state error covariance eigendistribution for the Riccati recursion of the RLS filter with shift structured regressors (21) for $n = 30$, $r = 2.5$, and $\beta = 0.5$

We omit the full proof of this lemma and instead offer some intuition as to why one may expect the result to hold. Note that any orthogonal matrix Θ is determined by $\frac{n(n-1)}{2}$ parameters, whereas a vector h_i is determined by only n parameters. Then it is plausible to think that the matrix Θ can be perturbed a *small* perturbation Δ so that $\Theta + \Delta$ remains orthogonal, yet it “rotates” h_i to lie along the first unit axis. The statement of the lemma gives a precise value on the “size” of Δ when Θ is generated randomly.

Going back to the proof of Theorem 1, we can see that all the steps were just algebraic manipulations except when we used the self-averaging property. This crucial step should be reexamined in the current scenario. In fact we can no longer assume that h_i and P_i are independent. However, using the aforementioned lemma, we can show that (21) still holds

$$\begin{aligned}
 & h_i(P_i - zI)^{-1}h_i^T \\
 &= h_i\Theta(\Theta P_i\Theta^T - zI)^{-1}\Theta^T h_i^T \\
 &= (\|h_i\|e_1 - h_i\Delta)(\bar{P}_i - zI)^{-1}(\|h_i\|e_1 - h_i\Delta)^T \\
 &= \|h_i\|^2 e_1(\bar{P}_i - zI)^{-1}e_1^T - 2\|h_i\|e_1(\bar{P}_i - zI)^{-1}\Delta^T h_i^T \\
 &\quad + h_i\Delta(\bar{P}_i - zI)^{-1}\Delta^T h_i^T \\
 &\rightarrow S(z) + 0 + 0
 \end{aligned} \tag{27}$$

Therefore the steady-state behavior will remain the same as the case of temporally white regressors.

Theorem 2: (RLS Riccati recursion with shift structured regressors): Consider the random Riccati recursion in (16) where $n \rightarrow \infty$ and $q = \frac{\beta}{n}$. Further assume that h_i have the shift structure described in (26). As $i \rightarrow \infty$, the Stieltjes transform of the eigendistribution of P_i converges to the steady-state Stieltjes transform, $S(z)$, which satisfies

$$\beta S(z) + c = -\log(r - z - z^2 S(z)). \tag{28}$$

where c is a constant which can be determined by insisting that the eigendistribution integrates to one.

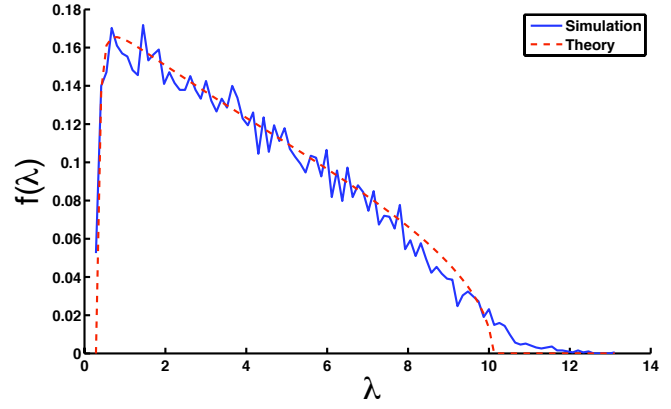


Fig. 4. Steady-state error covariance eigendistribution for the Riccati recursion of the RLS filter with intermittent observations (26) for $n = 30$, $\epsilon = 0.9$, $r = 2.5$, and $\beta = 0.5$

Proof: The proof follows that of Theorem 1. The only difference is that (27) should be used in order to ensure that the self-averaging property holds in (21). ■

We have to mention that the second term in the right-hand side behaves like $\frac{1}{\sqrt{n}}$. Therefore the convergence to the asymptotic result will be much slower in terms of n compared to the case of uncorrelated regressors. Figure 3 shows the simulation results for $n = 30$ and $\beta = 0.5$. We have also plotted the empirical distribution for the cases where the successive regressors differ by two and three shifts, respectively (in which case more randomness is brought to the regressor vectors at each time step). As expected, the empirical results converge much faster to the theoretical one due to more freedom in choosing Δ .

Remark So far we have only considered single measurements. The results can readily be extended to the case of multiple measurements as long as the number of measurements, m , is much less than the state vector size, n ($m \ll n$). When m is comparable to n , the results presented here are not valid. (A more complicated analysis is required which is the subject of another paper.)

V. RLS FILTER WITH INTERMITTENT OBSERVATIONS

A great deal of research has been devoted recently to the study of estimation and control over lossy networks [14], [15]. In this section we show how the problem of RLS filtering with intermittent observations can be handled in our framework. Once again consider the RLS filter (1) where the measurements y_i are lost with some known probability ϵ . It can be shown that [14] the error covariance undergoes both time and measurement update whenever a measurement is received,

$$P_{i+1} = P_i - \frac{P_i h_i^* h_i P_i}{r + h_i P_i h_i^*} + \frac{\beta}{n} I \quad \text{with prob. } 1 - \epsilon, \tag{29}$$

otherwise it undergoes only a time update

$$P_{i+1} = P_i + \frac{\beta}{n} I \quad \text{with prob. } \epsilon. \tag{30}$$

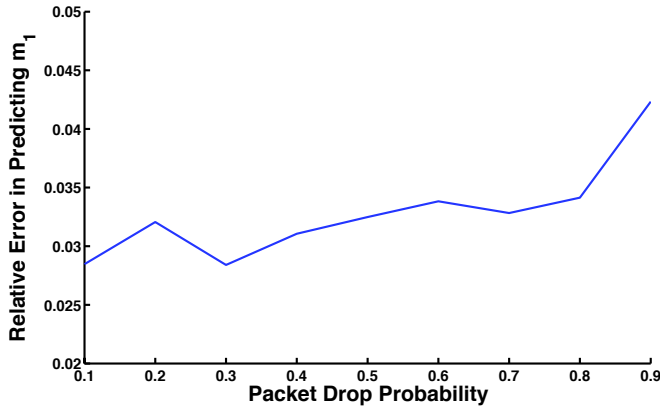


Fig. 5. Relative error in predicting $m_1 = \mathbb{E}\lambda = \frac{1}{n}\mathbb{E}\text{tr}P$ for an RLS filter with intermittent observations for different values of packet drop probability with $n = 20$, $r = 2.5$, and $\beta = 2.5$

We have assumed that the variance of measurement noise scales like $\frac{1}{n}$ following the discussion in Section III. Now one can combine the two recursions above to get a single Riccati recursion

$$P_{i+1} = P_i - \frac{P_i h_i^* h_i P_i}{r_i + h_i P_i h_i^*} + \frac{\beta}{n} I \quad (31)$$

where $\{r_i\}$ is now a collection of independent random variables such that,

$$r_i = \begin{cases} \infty & \text{with probability } \epsilon \\ r & \text{with probability } 1 - \epsilon \end{cases}.$$

The following theorem holds,

Theorem 3 (RLS Filter with Intermittent Observations):

For the RLS adaptive filter described above the error covariance undergoes the random Riccati recursion as in (31). One can show that the eigendistribution does converge for $\epsilon < 1$ and β being a finite constant. When $n \rightarrow \infty$, the steady-state Stieltjes transform satisfies

$$\beta S(z) + c = -(1 - \epsilon) \log(r - z + z^2 S(z)), \quad (32)$$

where c is a constant which is determined by insisting that the inverse Stieltjes transform should integrate to one.

Proof: We first write the Stieltjes transform at instance $i + 1$ as the average of the Stieltjes transforms conditioned on whether a packet drop has occurred or not,

$$\begin{aligned} S_{i+1}(z + \frac{\beta}{n}) &= S_{i+1}(z + \frac{\beta}{n})|_{r_i=\infty} p(r_i = \infty) \\ &\quad + S_{i+1}(z + \frac{\beta}{n})|_{r_i=r} p(r_i = r) \end{aligned} \quad (33)$$

Therefore,

$$\begin{aligned} S_{i+1}(z + q) &= -\frac{1}{n} \mathbb{E} \frac{d}{dz} \log \det(P_i - zI) - \frac{1 - \epsilon}{n} \mathbb{E} \frac{d}{dz} \\ &\quad \log \det \left(I - (P_i - zI)^{-1} \frac{P_i h_i^T h_i P_i}{r + h_i P_i h_i^T} \right). \end{aligned} \quad (34)$$

The rest of the proof follows the proof of Theorem 1. ■

Once again the distribution as well as the constant c should be determined numerically. We have done so in Figure 4 to obtain the theoretical prediction curve. It can be seen that our result closely predicts the empirical eigenvalue distribution. The relative error between the numerical value of $m_1 = \mathbb{E}\lambda = \frac{1}{n}\mathbb{E}\text{tr}P$ found through the above theorem and the one resulted from Monte Carlo simulation of an RLS filter with intermittent observations is plotted in Figure 5 for different values of packet drop probability and $n = 20$.

VI. CONCLUSIONS

In this work we used ideas from the theory of large random matrices and the Stieltjes transform to determine the eigenvalue distribution of the error covariance matrix for an RLS adaptive filter in three different scenarios. Namely temporally uncorrelated regressors, shift structured regressors, and the RLS filter with intermittent observation. We established a framework to find expressions which in many cases have appeared for the first time in the literature. This work is also a contribution to the random matrix theory in the sense that the form of Stieltjes transforms we have obtained have never been encountered in random matrix theory before (to the best of our knowledge).

REFERENCES

- [1] E. Eleftheriou and D. D. Falconer, "Tracking properties and steady-state performance of RLS adaptive filter algorithms," *IEEE Trans. Acoust. Speech Signal Process.*, vol. 34, pp. 1097–1110, 1986.
- [2] A.H. Sayed, *Fundamentals of Adaptive Filtering*, Wiley-IEEE, 2003.
- [3] S. Haykin, *Adaptive Filter Theory*, Prentice-Hall, 2001.
- [4] E. Eweda, "Comparison of RLS, LMS, and sign algorithms for tracking randomly time-varying channels," *IEEE Transactions on Signal Processing*, vol. 42, no. 11, pp. 2937–2944, November 1994.
- [5] B. Hassibi, A.H. Sayed, and T. Kailath, *Indefinite Quadratic Estimation and Control: A Unified Approach to H^2 and H^∞ Theories*, SIAM, 1999.
- [6] T. Kailath, A.H. Sayed, and B. Hassibi, *Linear Estimation*, Prentice-Hall, 2000.
- [7] E. Wigner, "Statistical properties of real symmetric matrices with many dimensions," in *Proc. 4th Canadian Mathematical Congress*, 1959, pp. 174–176.
- [8] E. Wigner, "On the distribution of roots of certain symmetric matrices," *Annals of Mathematics*, vol. 67, pp. 325–327, 1958.
- [9] A. M. Tulino and S. Verdú, "Random matrix theory and wireless communications," *Foundations and Trends in communications and Information Theory*, vol. 1, no. 1, pp. 1–182, 2004.
- [10] V. A. Marcenko and L. A. Pastur, "Distributions of eigenvalues for some sets of random matrices," *Math. USSR-Sbornik*, vol. 1, pp. 457–483, 1967.
- [11] J. Silverstein and S. Choi, "Analysis of the limiting spectral distribution of large dimensional random matrices," *Journal of Multivariate Analysis*, vol. 54, no. 2, pp. 295–309, 1995.
- [12] Z. D. Bai and J. W. Silverstein, "No eigenvalues outside the support of the limiting spectral distribution of large dimensional sample covariance matrices," *Annals of Probability*, vol. 26, pp. 316–345, 1998.
- [13] A. Vakili and B. Hassibi, "On the asymptotic eigenvalue distribution of certain random Lyapunov and Riccati recursions," in *Mathematical Theory of Networks and Systems*, 2008.
- [14] B. Sinopoli, L. Schenato, M. Franceschetti, K. Poolol, M. Jordan, and S. Sastry, "Kalman filtering with intermittent observations," *IEEE Transactions on Automatic Control*, vol. 49, no. 9, pp. 1453–1464, Sep 2004.
- [15] V. Gupta, A. F. Dana, J. Hespanha, R. M. Murray, and B. Hassibi, "Data transmission over networks for estimation and control," *IEEE Transactions on Automatic Control*, submitted 2006.